

DNA STATISTICAL ANALYSIS FOR Y-STR DATA

A. SCOPE

Y-STR loci are located on the non-recombining part of the Y chromosome. The loci should be considered linked and a database consisting of haplotypes must be used to generate statistics. The U.S. Y-STR database will be used.

A.1 COUNTING METHOD

The basis for the haplotype frequency estimation is the counting method. The number of times the matching haplotype has been observed in the combined data from all populations (African American, Asian, Caucasian, Hispanic, and Native American) in the U.S. Y-STR Database will be determined. The application of a confidence interval corrects for database size and sampling variation.

A.2 CONFIDENCE LIMIT

The 95% confidence limit shall be applied to the # of observations of the haplotype in the database.

- A.2.1 If the haplotype has not been previously observed in the database, the formula used to calculate the upper 95% confidence interval is:

$$1-(0.05)^{1/n}$$

where n is the size of the database

- A.2.2 If the haplotype has been observed in the database, the formula used for calculating the 95% confidence limit is that of an exact confidence interval (Clopper and Pearson). Let the number of samples in the database be n and the number of matches x . This gives a sample proportion p . Clopper and Pearson seeks a value p , such that

$$\sum_{k=0}^x \binom{n}{k} p_0^k (1-p_0)^{n-k} = 0.05$$

A.3 MIXTURE CALCULATIONS

The number of times a discernable dominant or minor haplotype has been observed in the U.S. Y-STR Database may be reported in the frequency format described above. When a mixture cannot

Document ID	Revision	Approval	Date Published
1777	7	Supervising Criminalist - Biology	12/29/2017 9:28:26 AM

be discerned, individuals may be excluded, however no inclusive conclusions can be offered because the Washoe County Sheriff's Office Crime Lab does not have a validated statistical application.

B. REFERENCES

- B.1 Forensic DNA Typing, John M. Butler, Academic Press 2001, 2005, and 2009
- B.2 WCSO validation of Yfiler
- B.3 Y-chromosome Short Tandem Repeat (Y-STR) Interpretation Guidelines Scientific Working Group on DNA Analysis Methods (SWGDAM), Forensic Science Communications, January 2009 - Volume 11 - Number 1
- B.4 The use of confidence or fiducial intervals illustrated in the case of the binomial. Biometrika (1934). **26**: p. 404-413. Clopper, C.J. and E.S. Pearson
- B.5 <http://www.usystrdatabase.org/>

Document ID	Revision	Approval	Date Published
1777	7	Supervising Criminalist - Biology	12/29/2017 9:28:26 AM